Statistique bayésienne en génétique des populations

Applications à l'étude des origines de l'homme

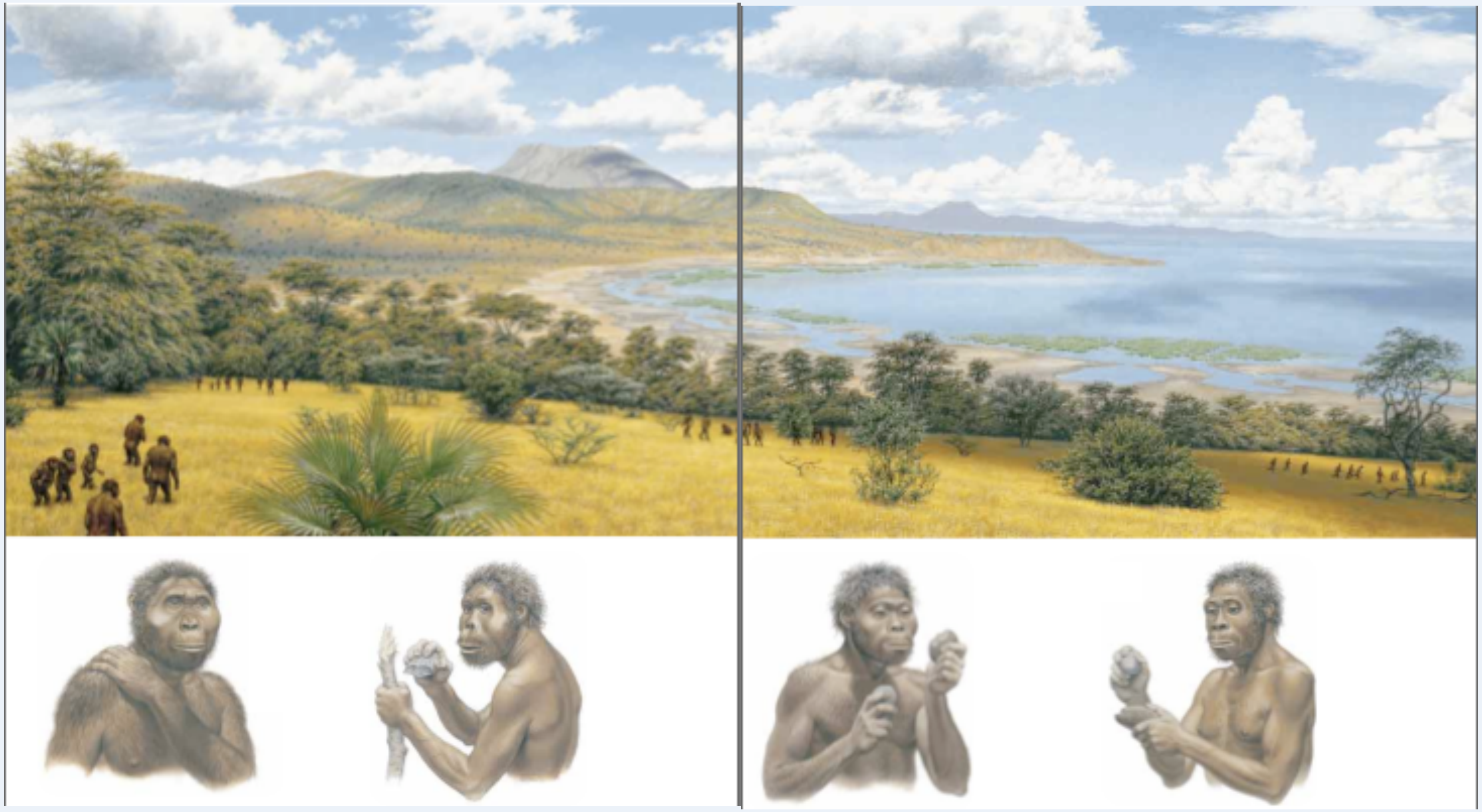
Michael Blum, TIMC-IMAG, Grenoble

Mattias Jakobsson, Uppsala University, Sweden

Rencontres Modélisation Mathématique et Biodiversité

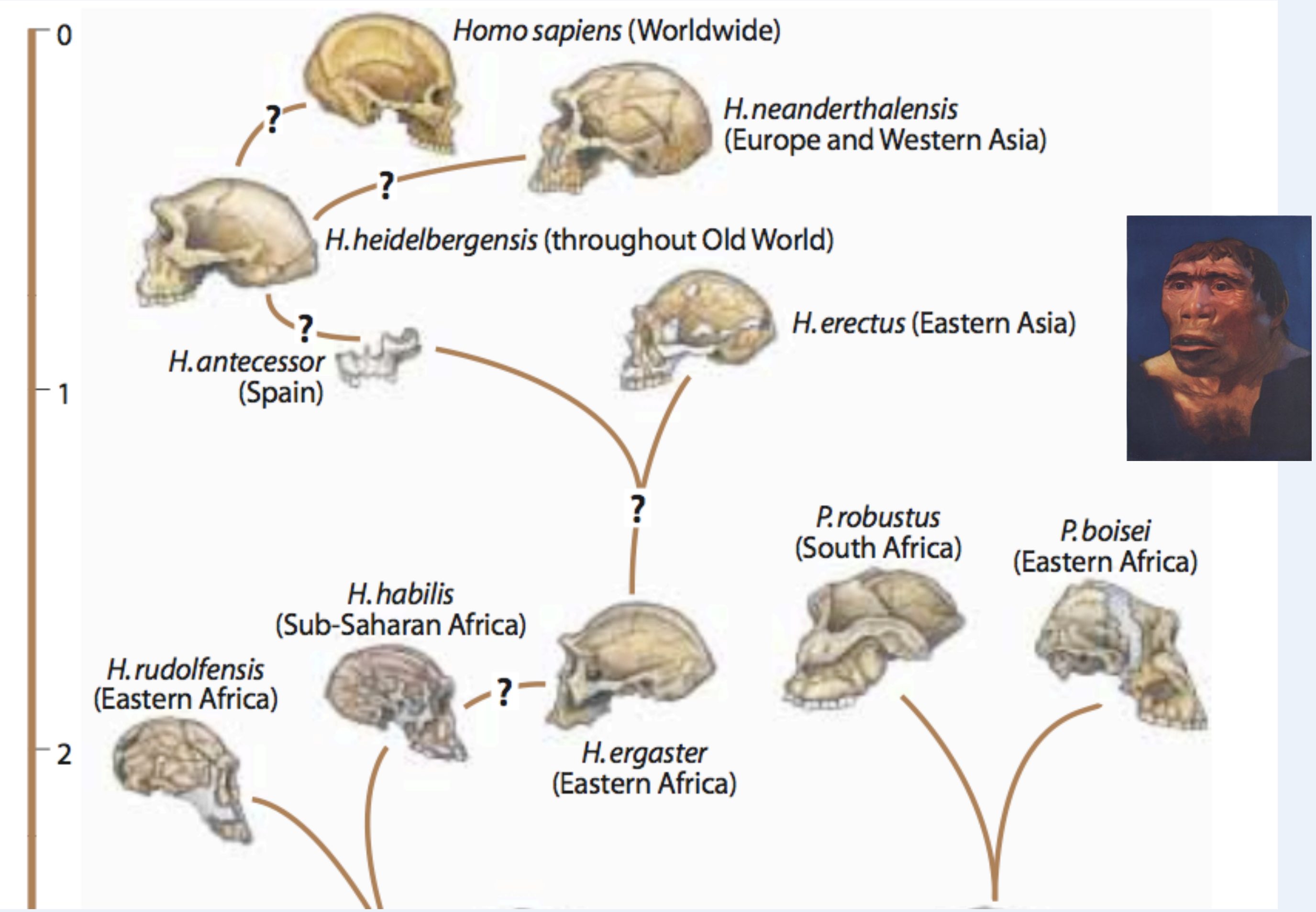
Ecole Polytechnique
21 Janvier 2010

Once we were not alone



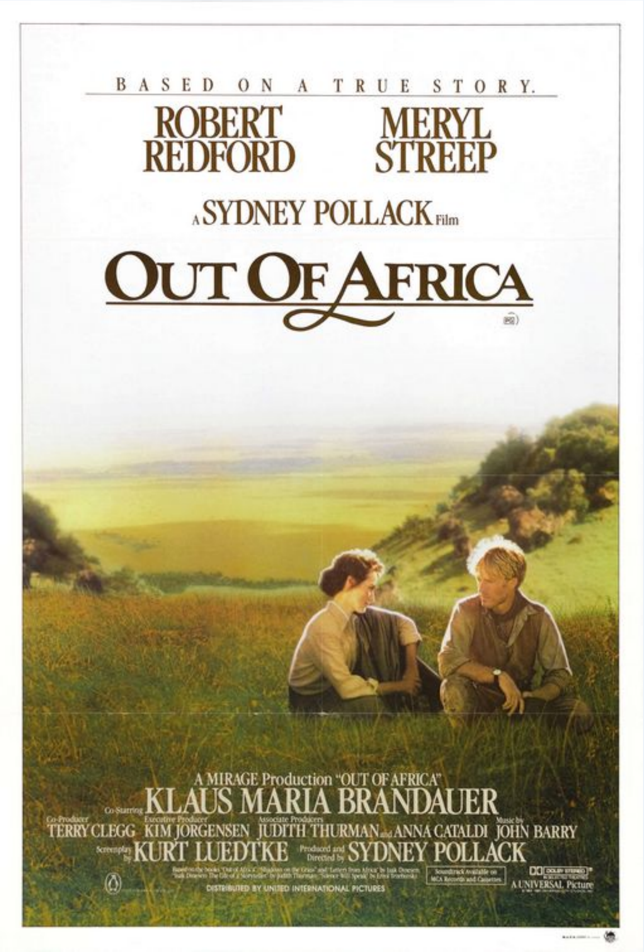
Tattersall, Scientific American, 1999

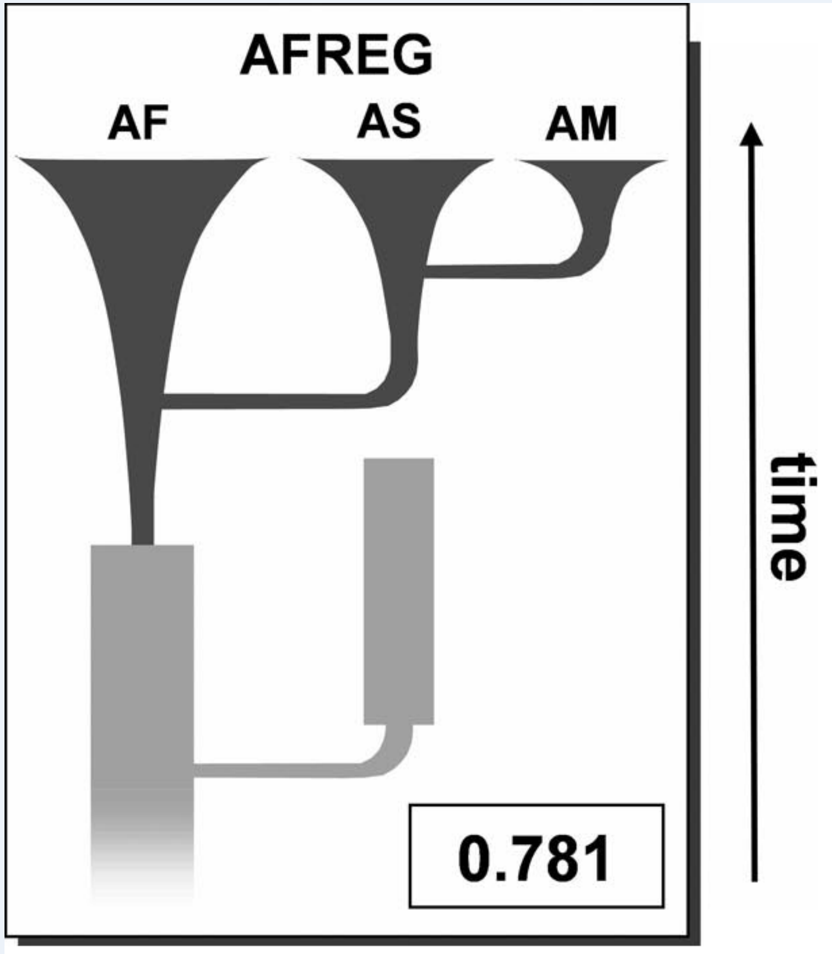
The roots of our solitude



Tattersall, Scientific American, 1999

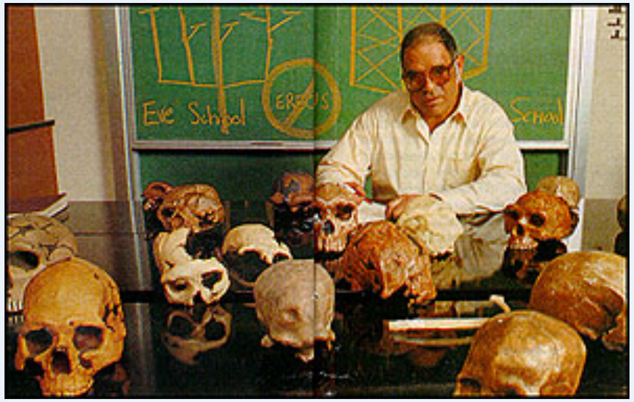
Out-Of-Africa

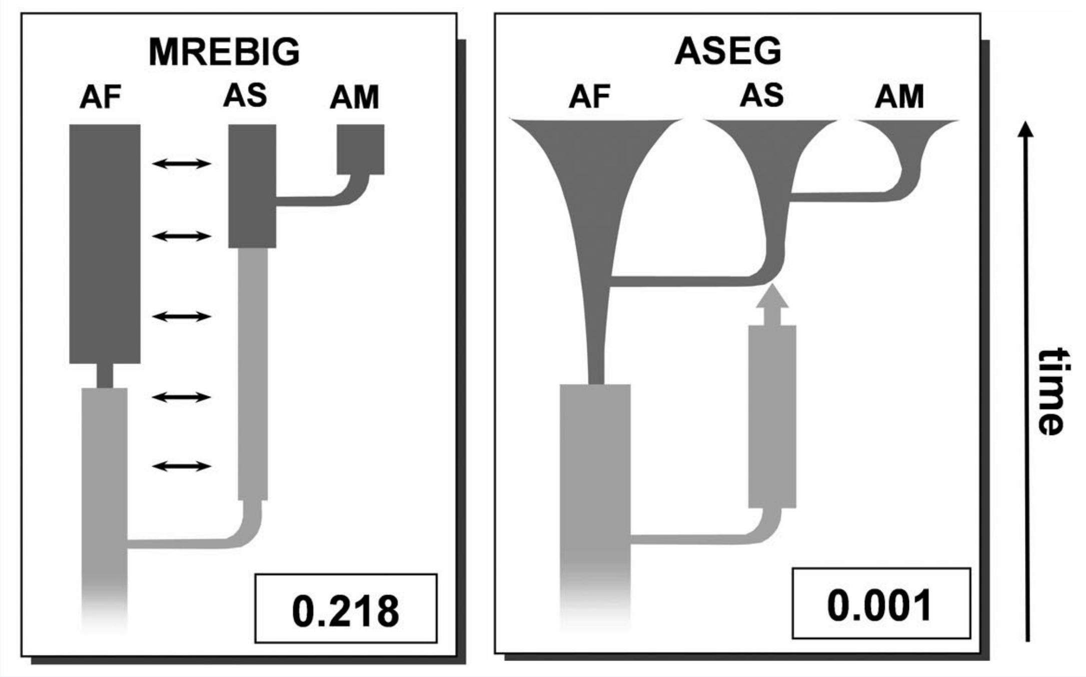




Fagundes et al. PNAS 2007

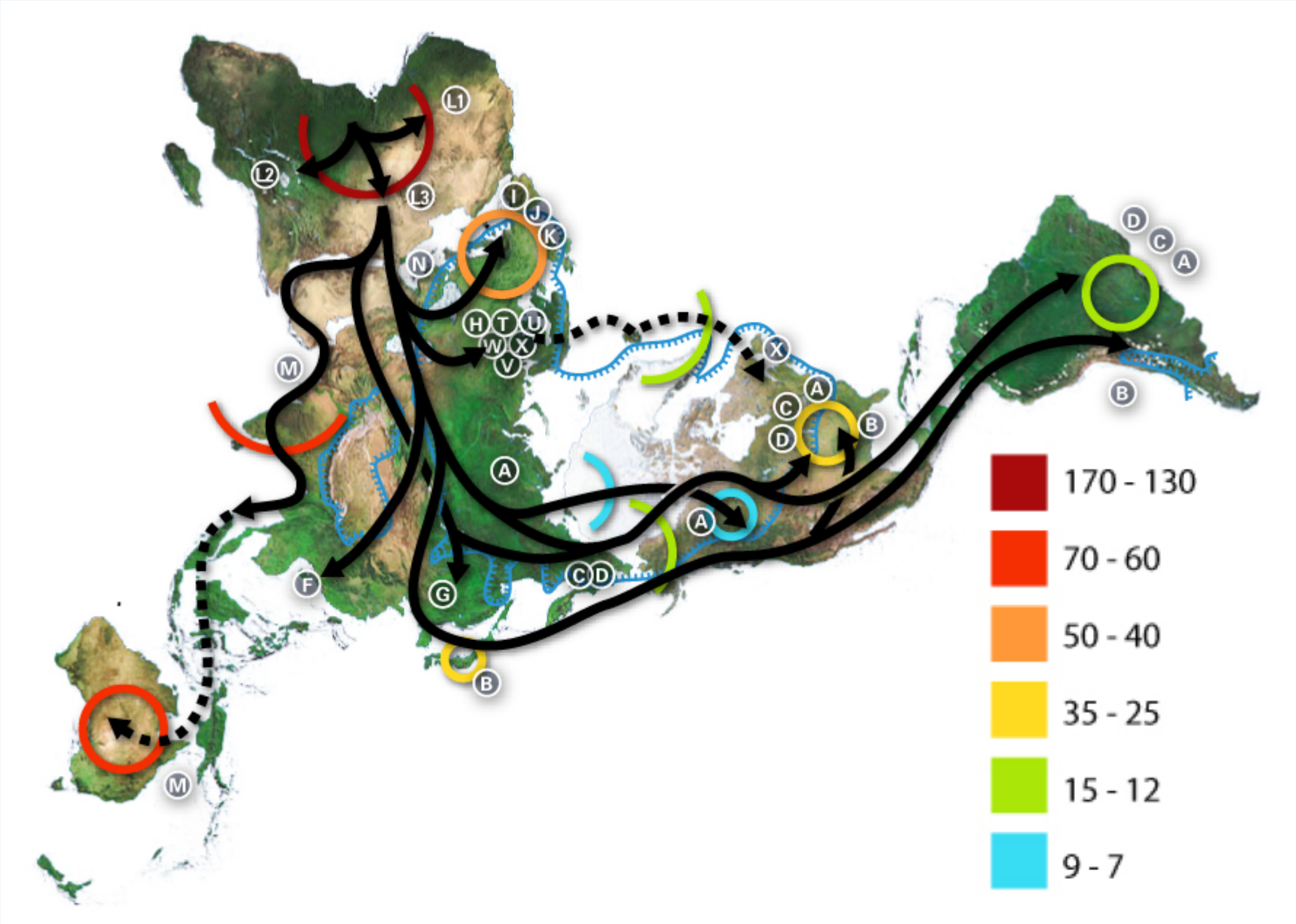
Multi-regional model





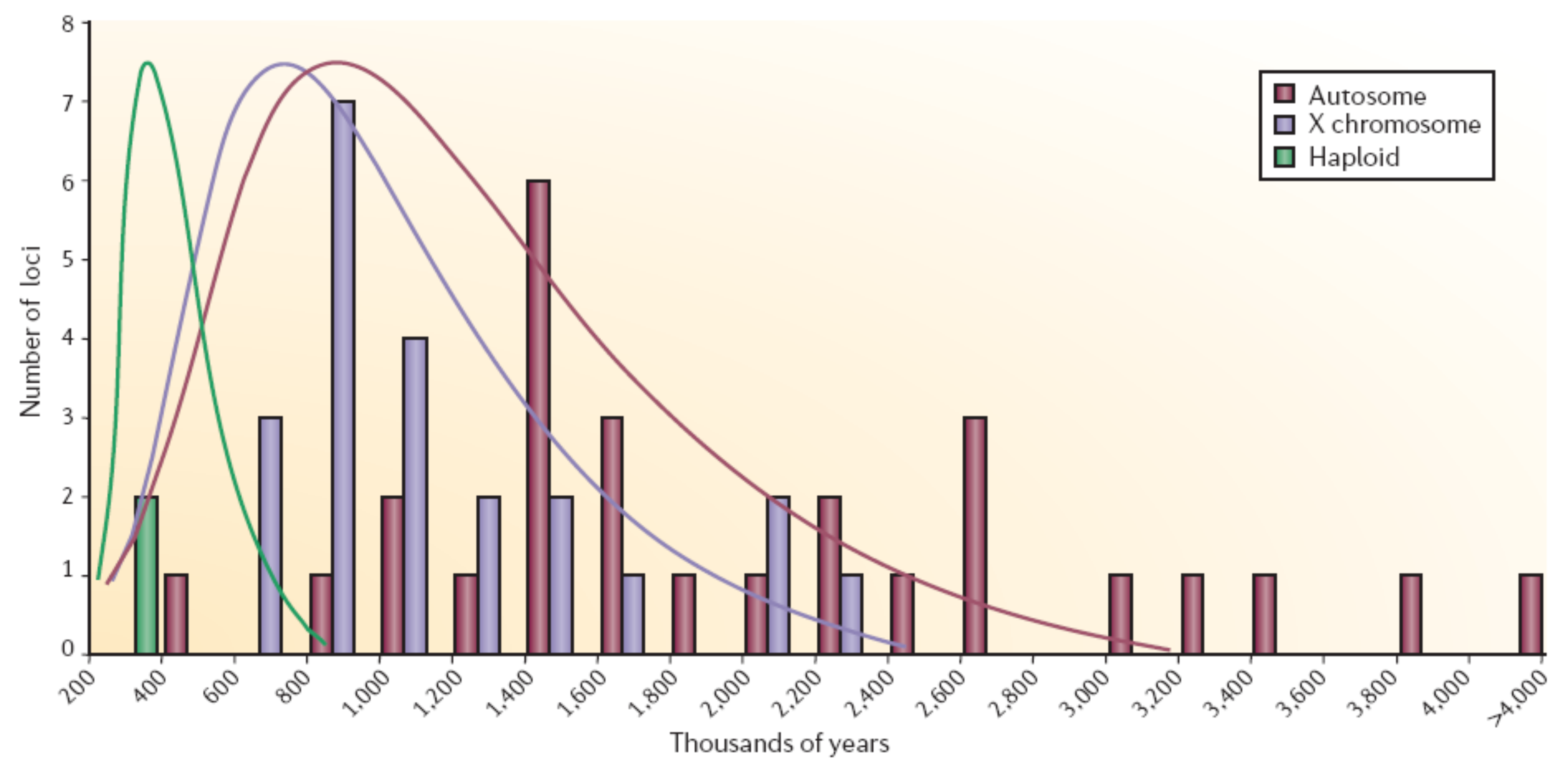
Fagundes et al. PNAS 2007

Mitochondrial TMRCA



Cann et al. Nature 1987

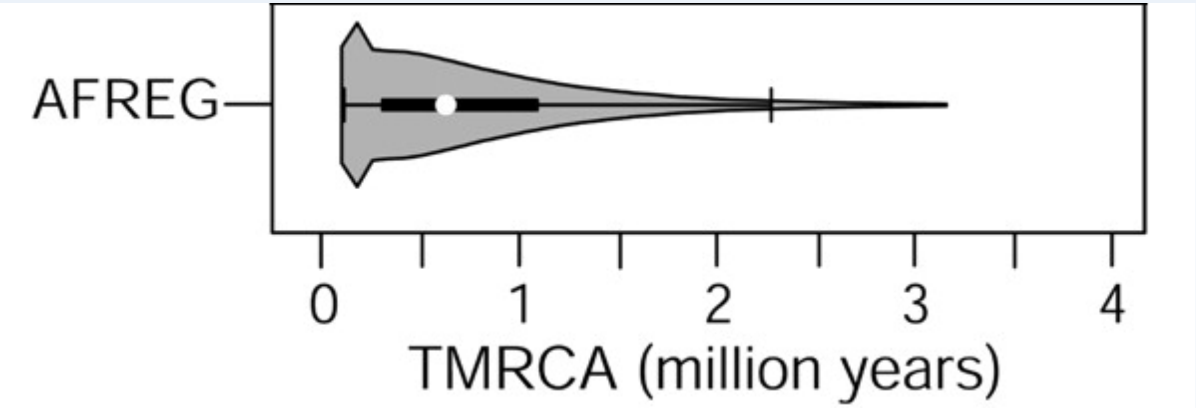
Autosomal TMRCA



Garrigan and Hammer Nat Rev Genet 2006

TMRCA in the Out Of Africa Model

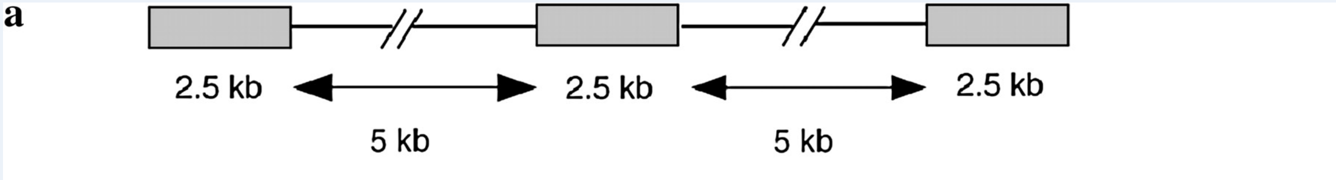
- Expected distribution

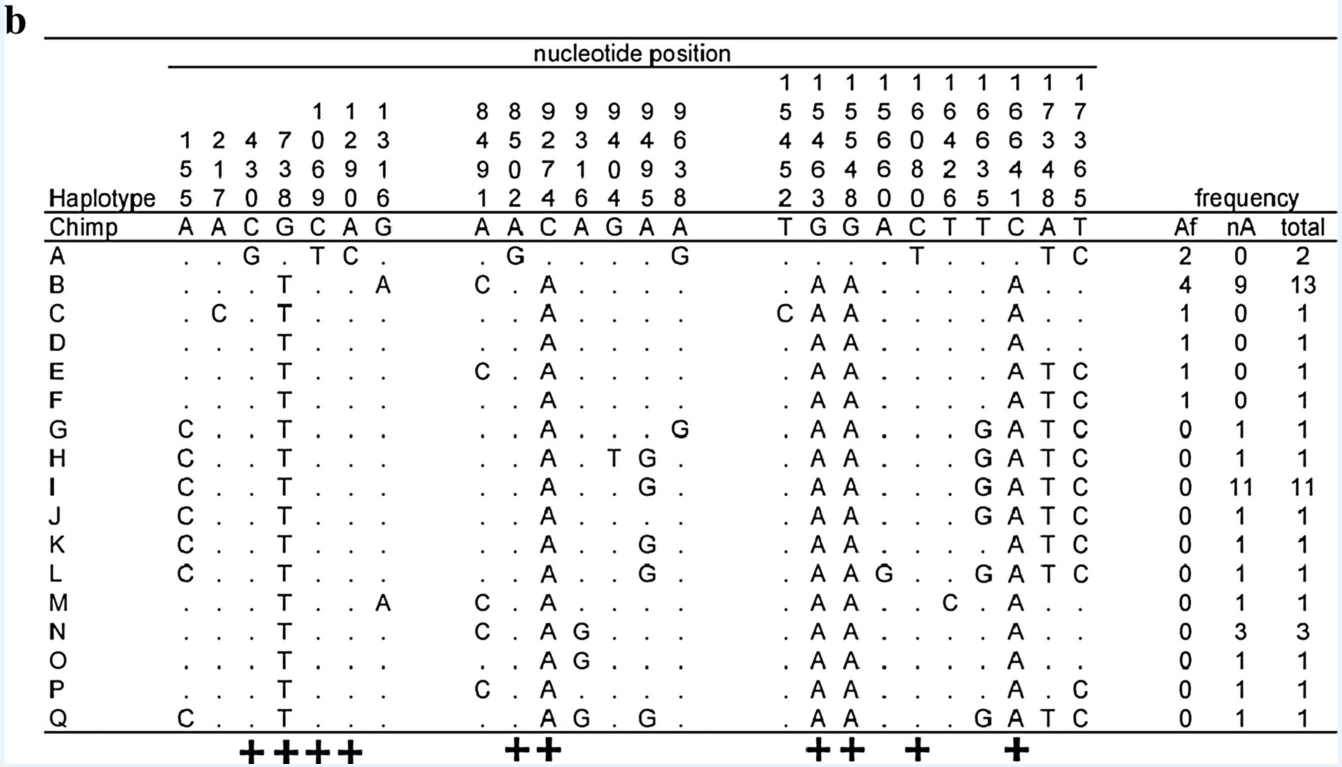


Fagundes et al. 2007

- “There might be too many loci with deep genealogical histories to be compatible with a simple, single origin model” Garrigan and Hammer 2006

Data

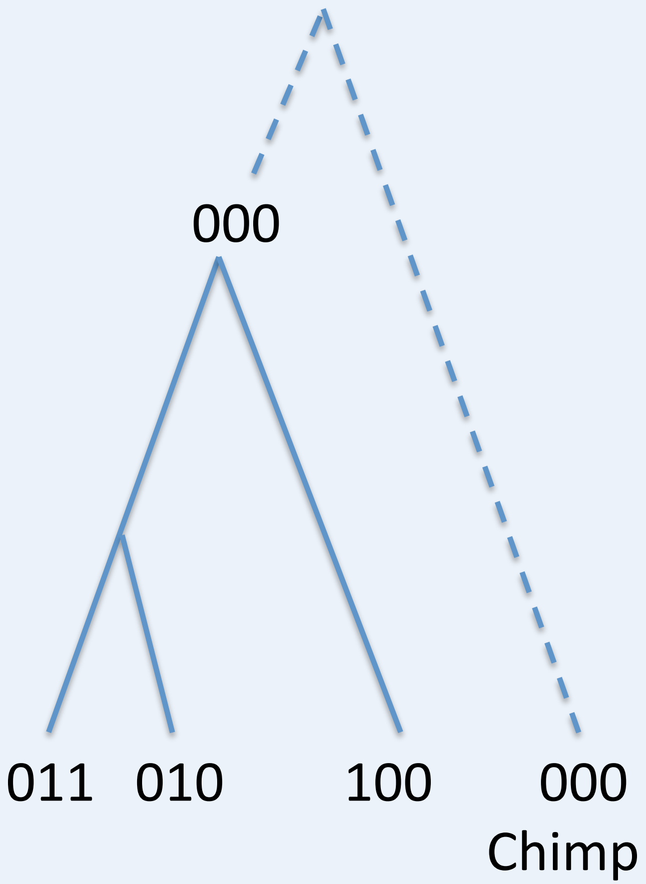




	nucleotide position																								frequency		
Haplotype	155	217	430	738	1069	1290	1316	8491	8502	9274	9316	9404	9495	9638	15452	15463	15548	15660	16080	16426	16635	16641	17348	17365	Af	nA	total
Chimp	A	A	C	G	C	A	G	A	A	C	A	G	A	A	T	G	G	A	C	T	T	C	A	T			
A	.	.	G	.	T	C	.	.	G	.	.	.	.	G	.	.	.	.	T	.	.	.	T	C	2	0	2
B	.	.	.	T	.	.	A	C	.	A	.	.	.	.	.	A	A	.	.	.	.	A	.	.	4	9	13
C	.	C	.	T	.	.	.	.	.	A	.	.	.	.	C	A	A	.	.	.	.	A	.	.	1	0	1
D	.	.	.	T	.	.	.	.	.	A	.	.	.	.	.	A	A	.	.	.	.	A	.	.	1	0	1
E	.	.	.	T	.	.	.	C	.	A	.	.	.	.	.	A	A	.	.	.	.	A	T	C	1	0	1
F	.	.	.	T	.	.	.	.	.	A	.	.	.	.	.	A	A	.	.	.	.	A	T	C	1	0	1
G	C	.	.	T	.	.	.	.	.	A	.	.	.	G	.	A	A	.	.	.	G	A	T	C	0	1	1
H	C	.	.	T	.	.	.	.	.	A	.	T	G	.	.	A	A	.	.	.	G	A	T	C	0	1	1
I	C	.	.	T	.	.	.	.	.	A	.	.	G	.	.	A	A	.	.	.	G	A	T	C	0	11	11
J	C	.	.	T	.	.	.	.	.	A	.	.	.	.	.	A	A	.	.	.	G	A	T	C	0	1	1
K	C	.	.	T	.	.	.	.	.	A	.	.	G	.	.	A	A	.	.	.	.	A	T	C	0	1	1
L	C	.	.	T	.	.	.	.	.	A	.	.	G	.	.	A	A	G	.	.	G	A	T	C	0	1	1
M	.	.	.	T	.	.	A	C	.	A	.	.	.	.	.	A	A	.	.	C	.	A	.	.	0	1	1
N	.	.	.	T	.	.	.	C	.	A	G	.	.	.	.	A	A	.	.	.	.	A	.	.	0	3	3
O	.	.	.	T	.	.	.	.	.	A	G	.	.	.	.	A	A	.	.	.	.	A	.	.	0	1	1
P	.	.	.	T	.	.	.	C	.	A	.	.	.	.	.	A	A	.	.	.	.	A	.	C	0	1	1
Q	C	.	.	T	.	.	.	.	.	A	G	.	G	.	.	A	A	.	.	.	G	A	T	C	0	1	1
			+	+	+	+			+	+						+	+		+			+					

HOMINID project
Wall et al., Mol Biol Evol 2008

TMRCA inference

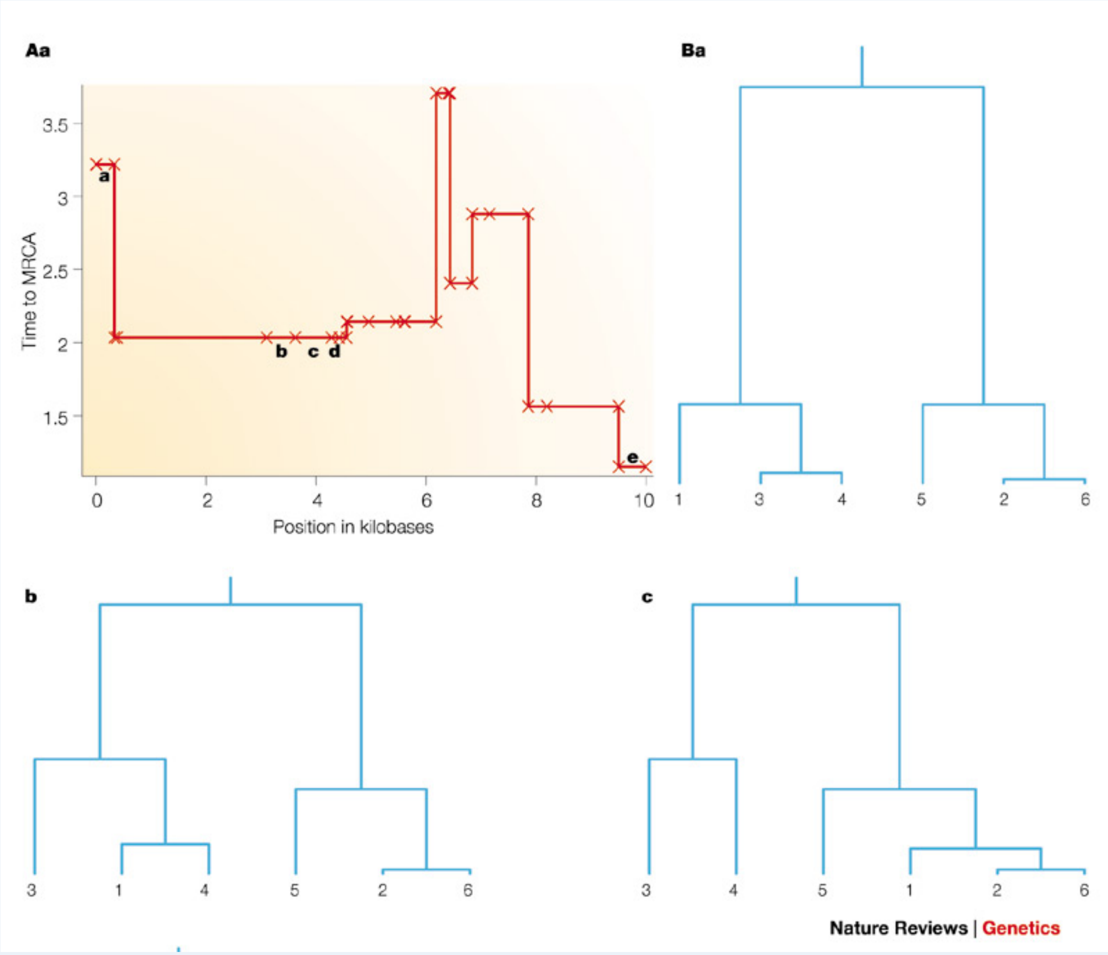


TMRCA = (2+1+1)/3 mutations

= 4/3 mutations

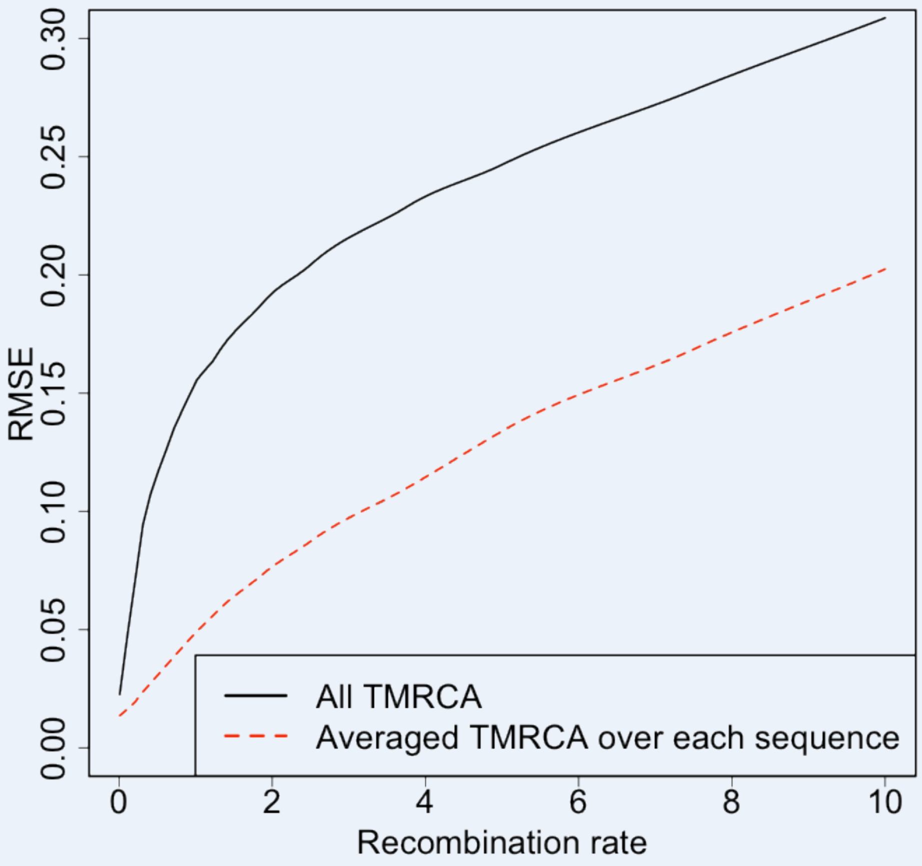
Thompson et al., PNAS 2000
Hudson, J Mol Evol 2007

TMRCA inference

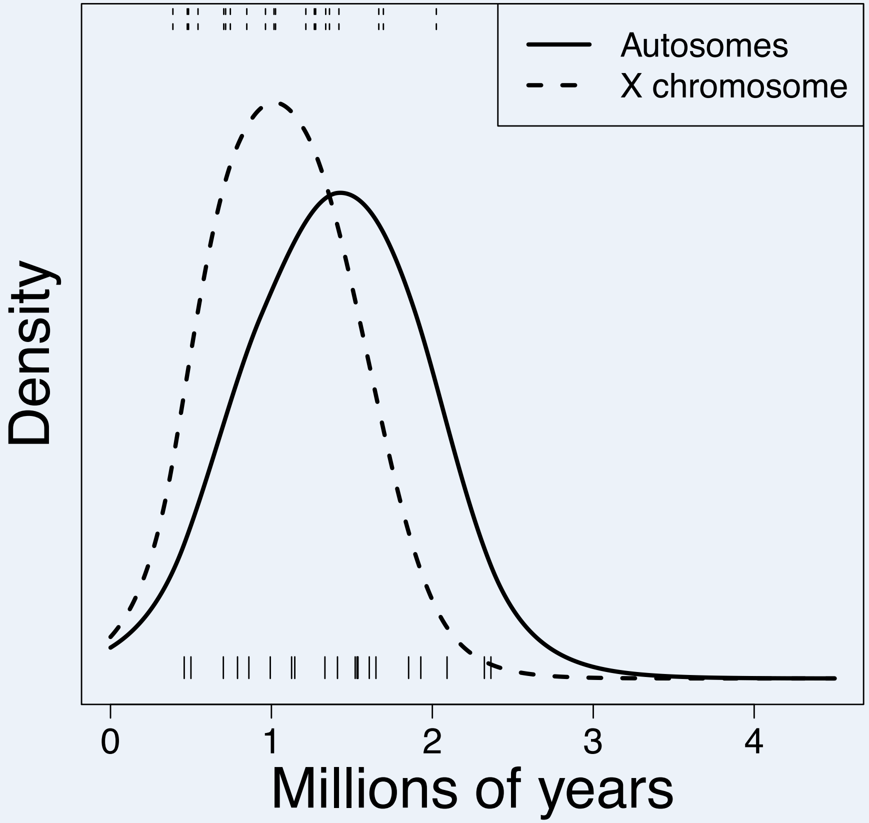


Rosenberg and Nordborg, Nature Rev Genet 2002

TMRCA inference

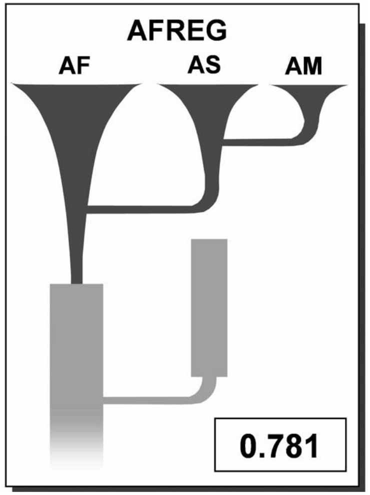


Genome-wide distribution of the TMRCA

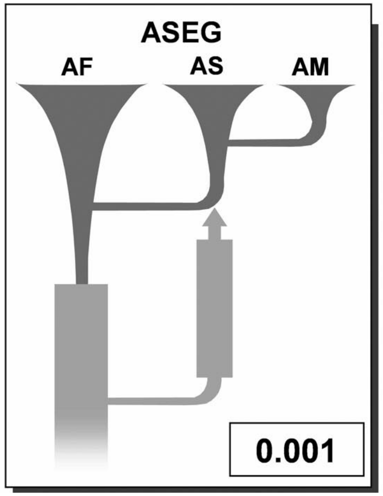


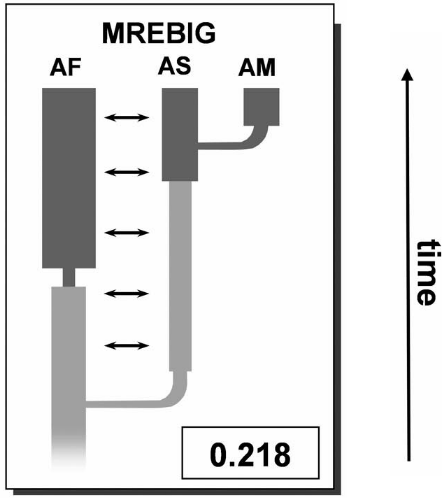
Models of human evolution

Out of Africa



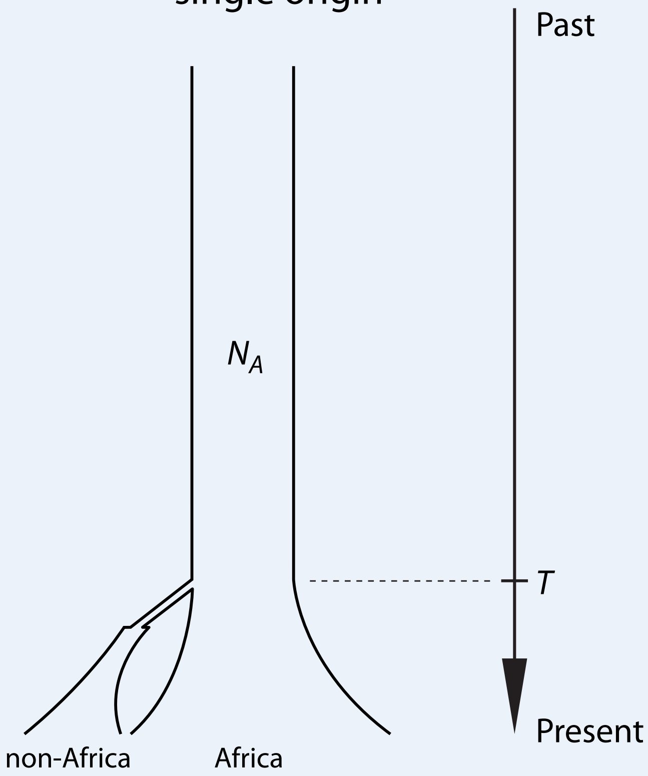
Multi-Regional



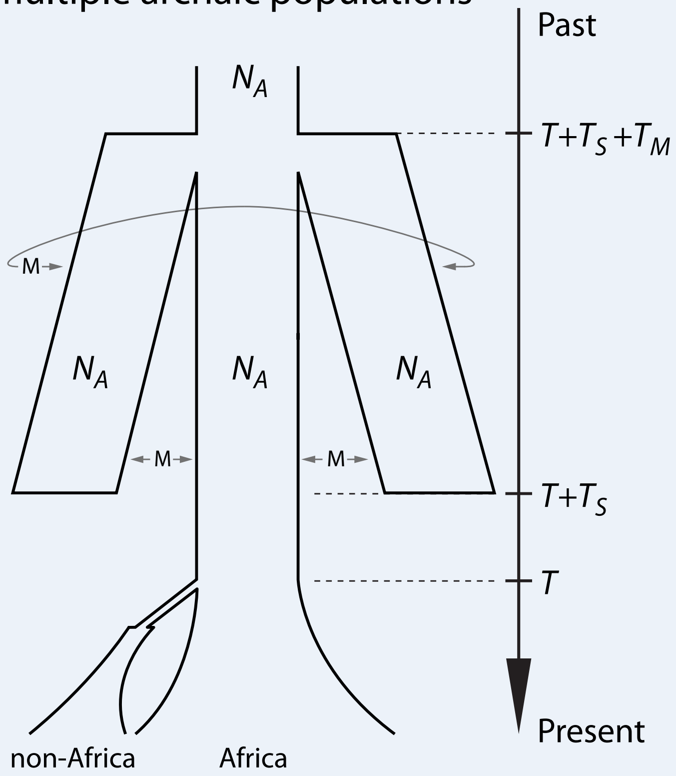


Models of human evolution
Archaic admixture in Africa

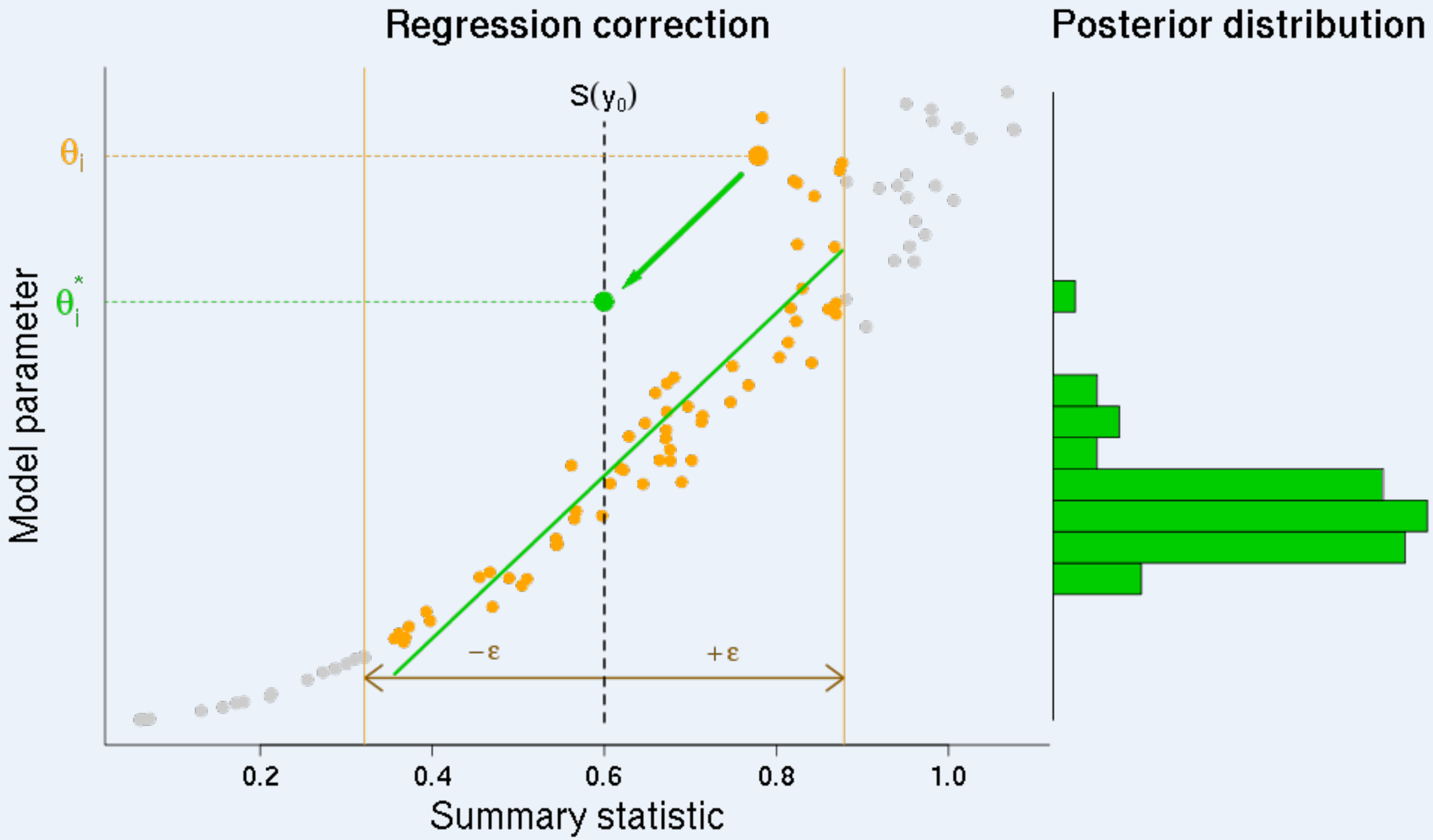
Out-of-Africa
single origin



Out-of-Africa
multiple archaic populations

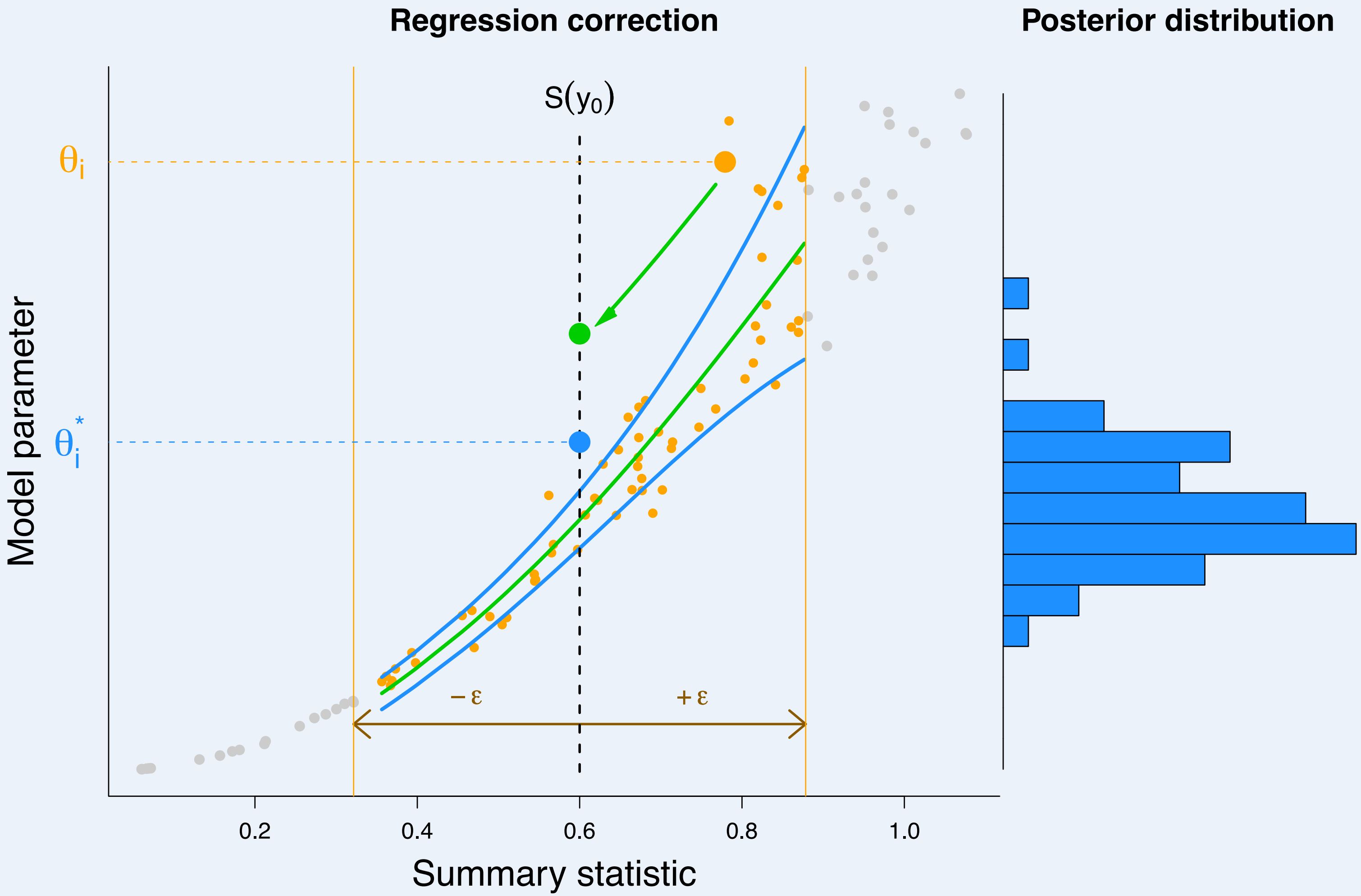


Approximate Bayesian Computation (ABC)
Parameter inference



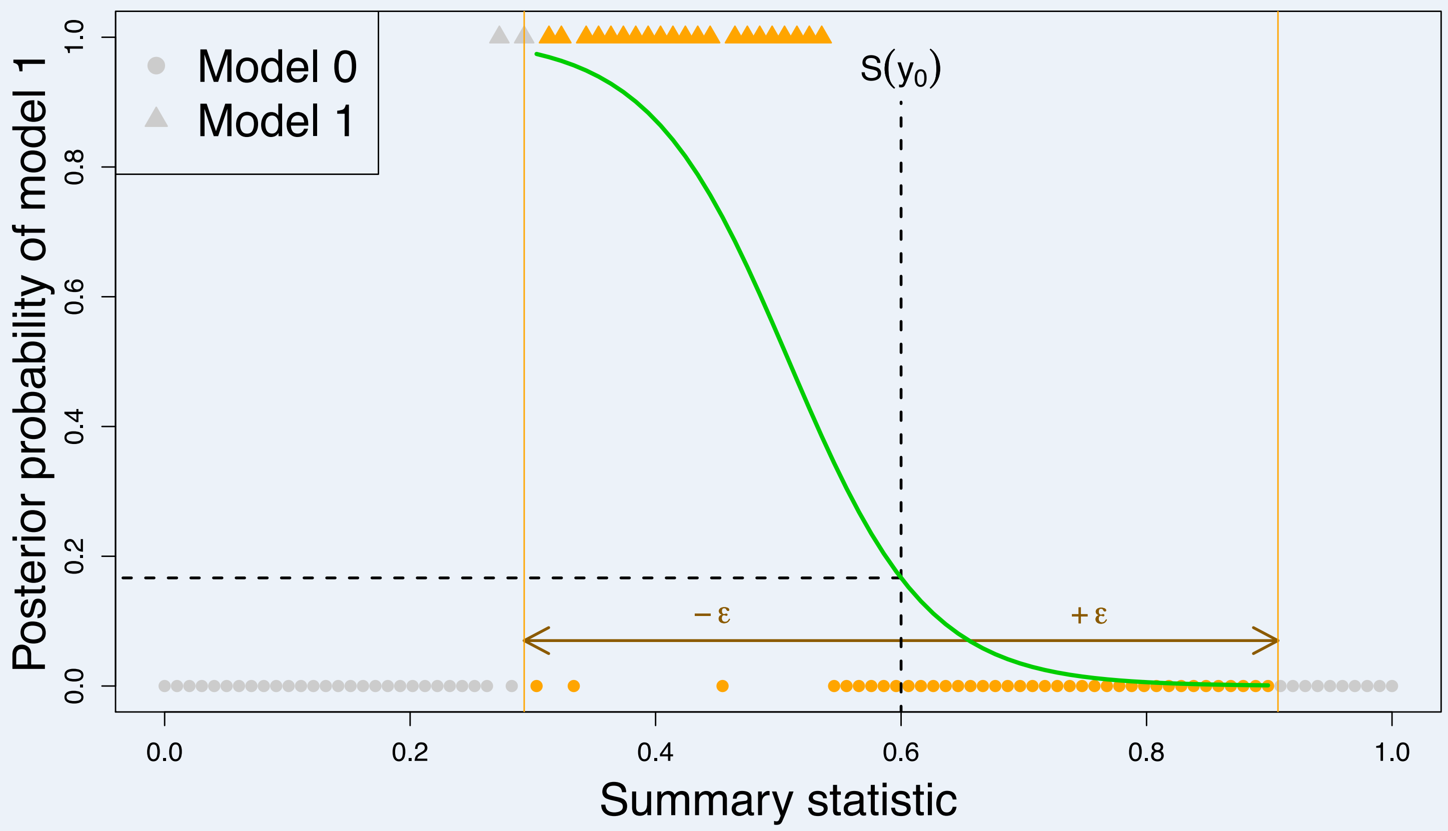
Beaumont et al., Genetics 2002

Approximate Bayesian Computation Parameter inference

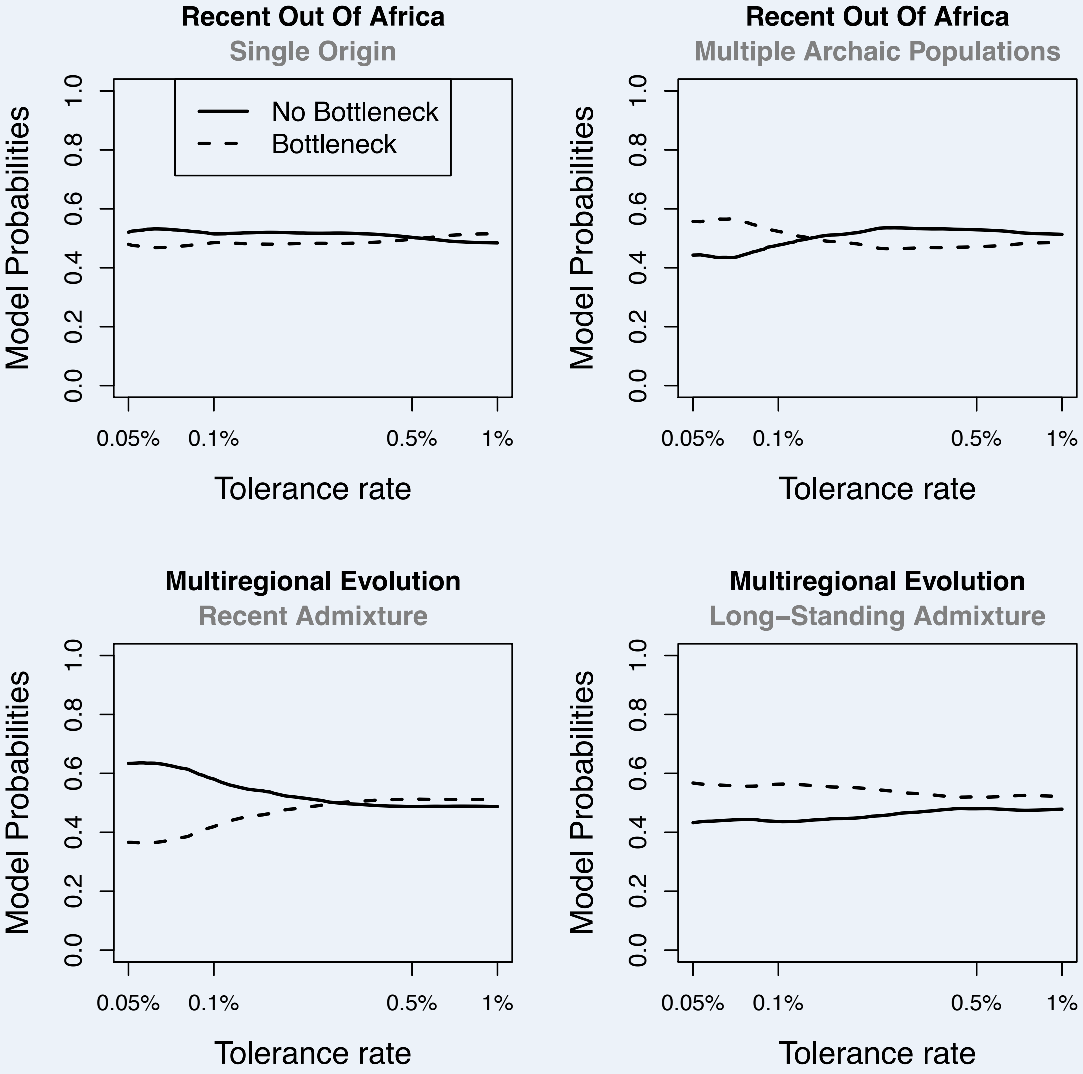


Blum and François, Statistics and Computing 2010

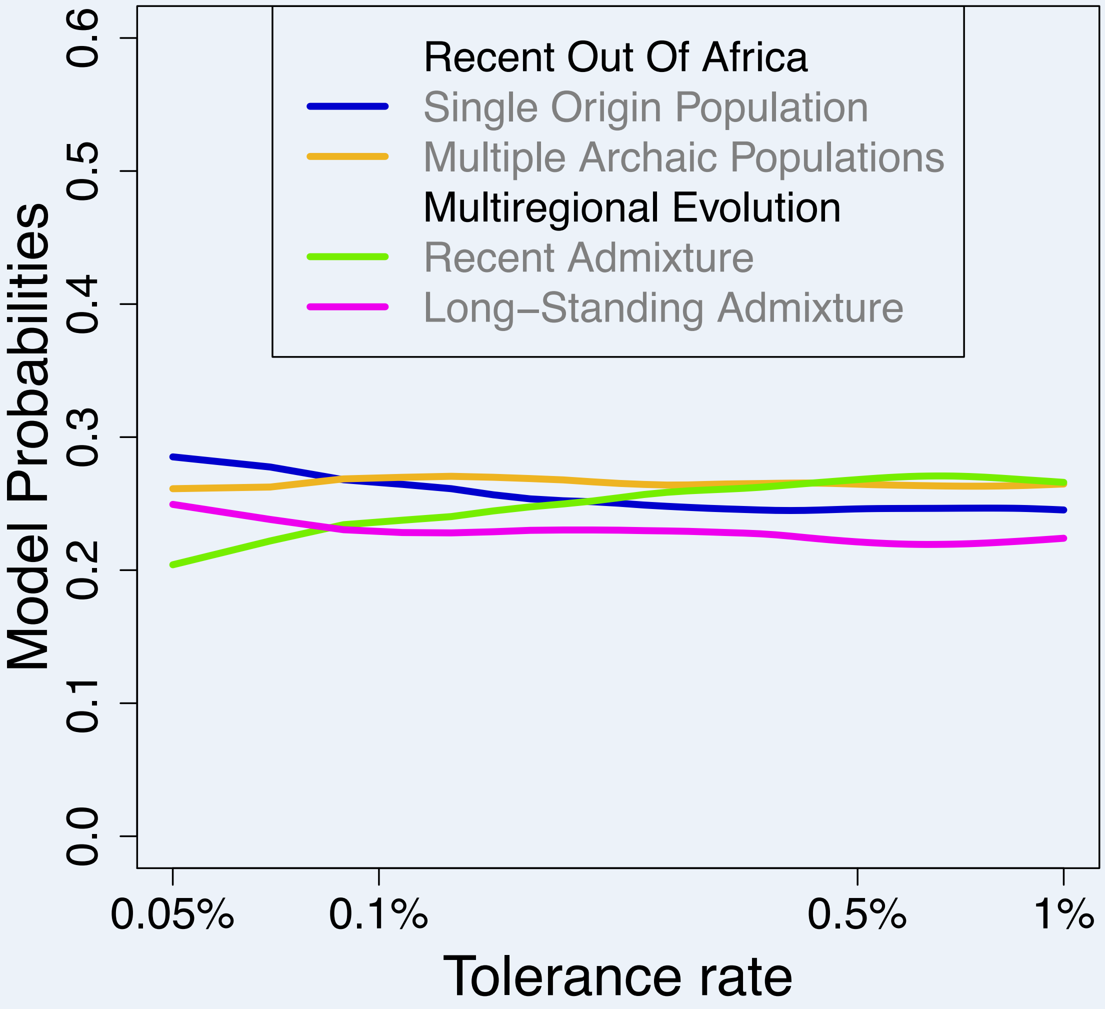
Approximate Bayesian Computation
Model selection



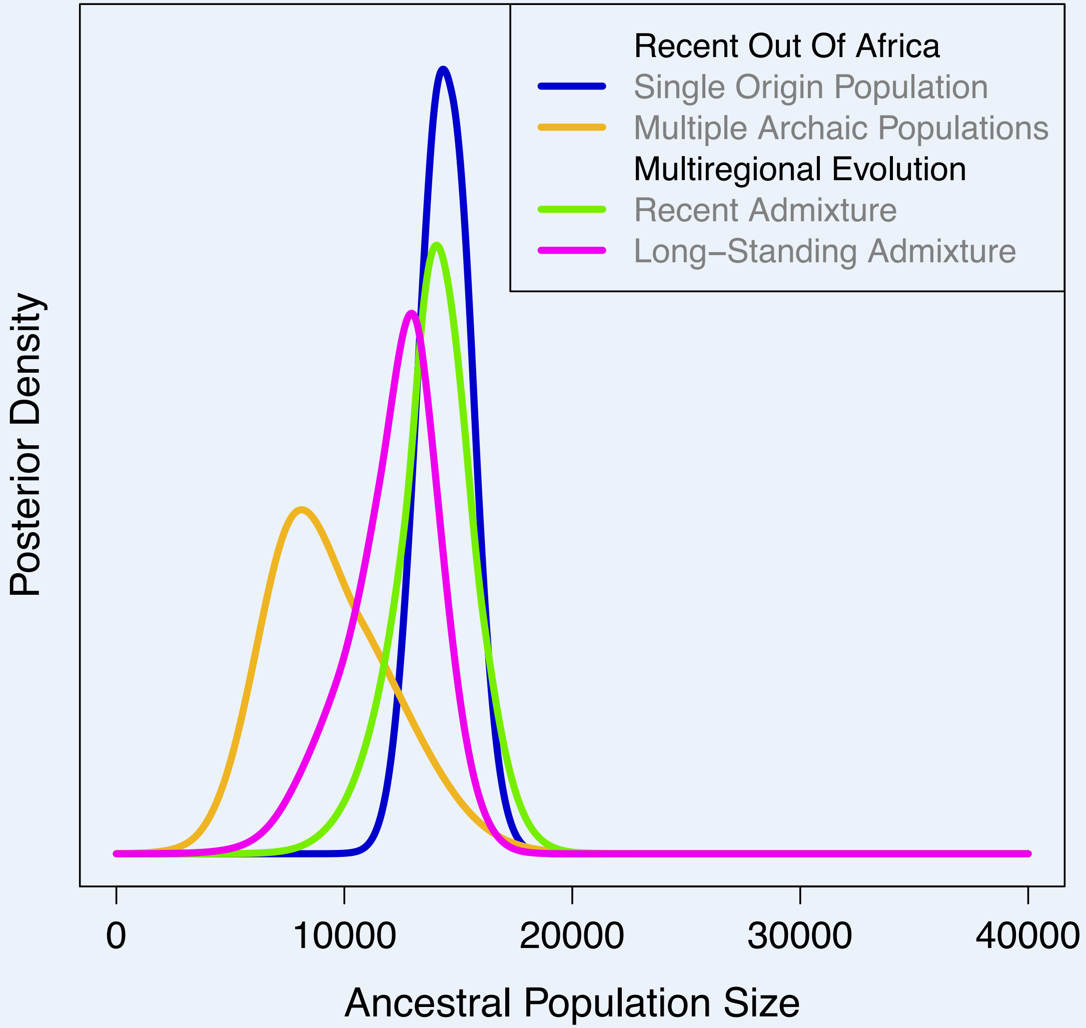
Testing the speciation bottleneck



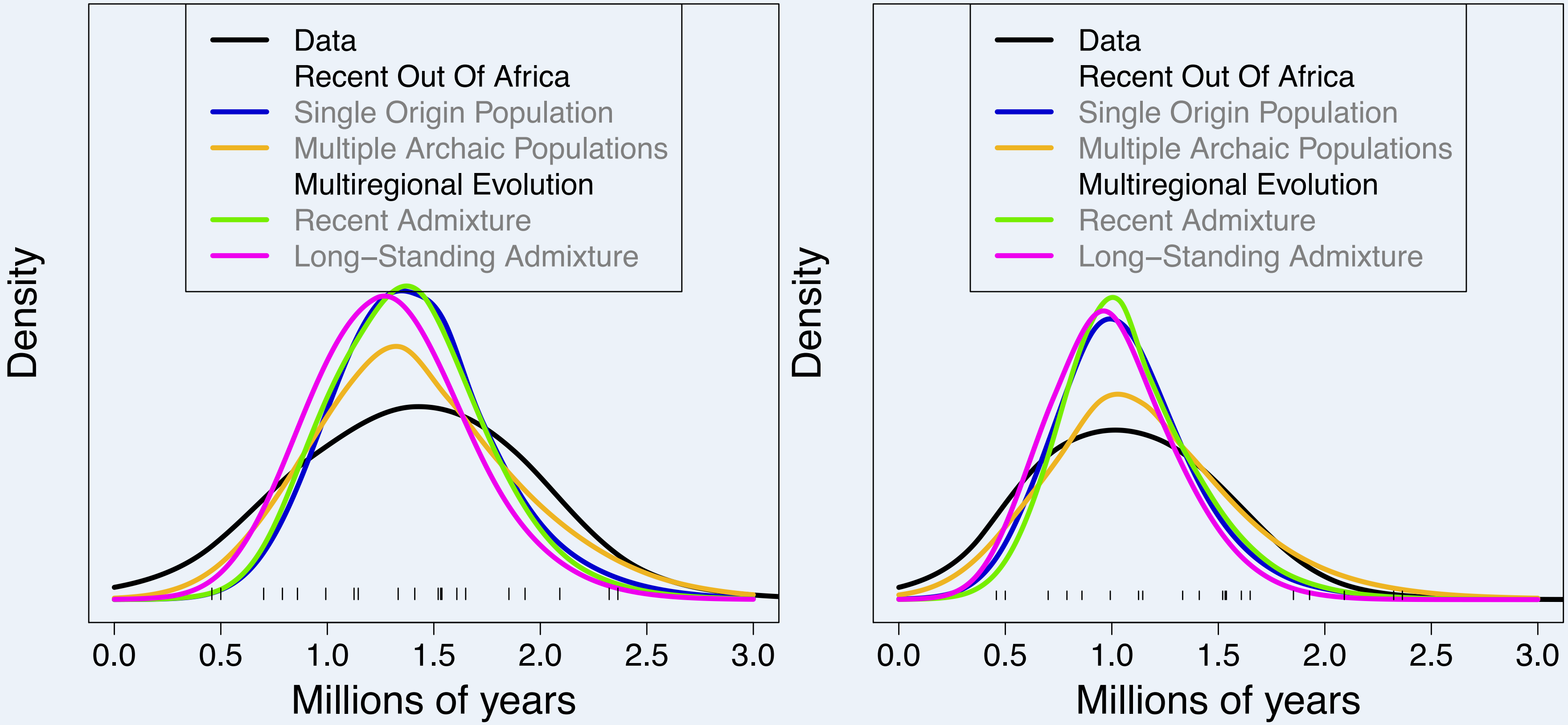
Archaic admixture VS Single Origin



Ancestral population size of humans



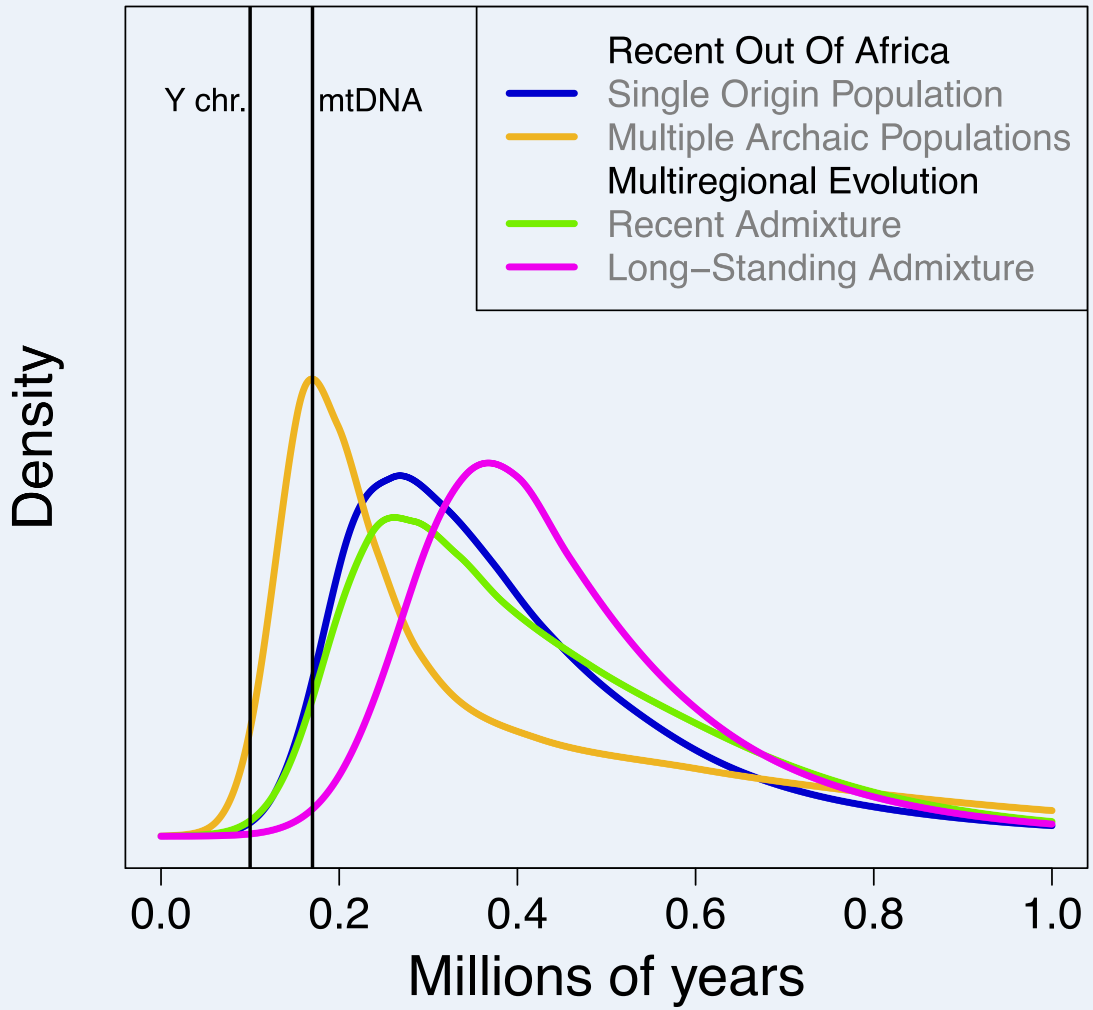
Goodness of fit



Autosomes

X chromosome

TMRCA for haploid sexual markers



Standard coalescent theory

- Out of Africa (~60,000 years)
- Autosomal TMRCA
  - 60,000+ 25*2*(2*14,000) ~ 1,400,000
- X TMRCA
  - 60,000+ 25*2*(3/2*14,000)~ 1,100,000
- mtDNA, Y chr.
  - 60,000+ 25*2*(14,000/2)~ 400,000

Conclusions

- Recent Out-of-Africa hypothesis
  - compatible with deep divergences of human gene trees
  - predicts smaller than “observed”-discrepancy between autosomal and haploid TMRCAs
- Archaic admixture in Africa
  - Large variance of TMRCAs
  - Large discrepancy between autosomal and haploid TMRCAs